

0590  
1128#2  
OIPE

## RAW SEQUENCE LISTING

DATE: 01/02/2002

PATENT APPLICATION: US/09/912,717

TIME: 14:52:54

Input Set : N:\Crf3\RULE60\09912717.raw

Output Set: N:\CRF3\01022002\I912717.raw

## SEQUENCE LISTING

## 3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Hillman, Jennifer L.

6 Corley, Neil C.

7 Baughn, Mariah R.

9 (ii) TITLE OF INVENTION: DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE  
10 HOMOLOG

12 (iii) NUMBER OF SEQUENCES: 3

14 (iv) CORRESPONDENCE ADDRESS:

15 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.

16 (B) STREET: 3174 Porter Drive

17 (C) CITY: Palo Alto

18 (D) STATE: CA

19 (E) COUNTRY: USA

20 (F) ZIP: 94304

ENTERED

22 (v) COMPUTER READABLE FORM:

23 (A) MEDIUM TYPE: Diskette

24 (B) COMPUTER: IBM Compatible

25 (C) OPERATING SYSTEM: Windows

26 (D) SOFTWARE: FastSEQ for Windows Version 2.0b

28 (vi) CURRENT APPLICATION DATA:

C--&gt; 29 (A) APPLICATION NUMBER: US/09/912,717

C--&gt; 30 (B) FILING DATE: 24-Jul-2001

31 (C) CLASSIFICATION:

33 (vii) PRIOR APPLICATION DATA:

34 (A) APPLICATION NUMBER: 09/565,910

35 (B) FILING DATE:

39 (viii) ATTORNEY/AGENT INFORMATION:

40 (A) NAME: Cerrone, Michael C

41 (B) REGISTRATION NUMBER: 39,132

42 (C) REFERENCE/DOCKET NUMBER: PF-0532 US

44 (ix) TELECOMMUNICATION INFORMATION:

45 (A) TELEPHONE: 650-855-0555

46 (B) TELEFAX: 650-855-0572

47 (C) TELEX:

50 (2) INFORMATION FOR SEQ ID NO: 1:

52 (i) SEQUENCE CHARACTERISTICS:

53 (A) LENGTH: 314 amino acids

54 (B) TYPE: amino acid

55 (C) STRANDEDNESS: single

56 (D) TOPOLOGY: linear

58 (vii) IMMEDIATE SOURCE:

59 (A) LIBRARY: PROSNON01

60 (B) CLONE: 2278458

62 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

64 Met Ser Val Gly Phe Ile Gly Ala Gly Gln Leu Ala Tyr Arg Phe Thr

65 1 5 10 15

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/912,717

DATE: 01/02/2002

TIME: 14:52:54

Input Set : N:\Crif3\RULE60\09912717.raw

Output Set: N:\CRF3\01022002\I912717.raw

```

66  Ala Ala Gly Ile Leu Ser Ala His Lys Ile Ile Ala Ser Ser Pro Glu
67          20          25          30
68  Met Asn Leu Pro Thr Val Ser Ala Leu Arg Lys Met Gly Val Asn Leu
69          35          40          45
70  Thr Arg Ser Asn Lys Glu Thr Val Lys His Ser Asp Val Leu Phe Leu
71          50          55          60
72  Ala Val Lys Pro His Ile Ile Pro Phe Ile Leu Asp Glu Ile Gly Ala
73          65          70          75          80
74  Asp Val Gln Ala Arg His Ile Val Val Ser Cys Ala Ala Gly Val Thr
75          85          90          95
76  Ile Ser Ser Val Glu Lys Lys Leu Met Ala Phe Gln Pro Ala Pro Lys
77          100         105         110
78  Val Ile Arg Cys Met Thr Asn Thr Pro Val Val Val Gln Glu Gly Ala
79          115         120         125
80  Thr Val Tyr Ala Thr Gly Thr His Ala Leu Val Glu Asp Gly Gln Leu
81          130         135         140
82  Leu Glu Gln Leu Met Ser Ser Val Gly Phe Cys Thr Glu Val Glu Glu
83          145         150         155         160
84  Asp Leu Ile Asp Ala Val Thr Gly Leu Ser Gly Ser Gly Pro Ala Tyr
85          165         170         175
86  Ala Phe Met Ala Leu Asp Ala Asp Gly Gly Val Lys Met Gly Leu Pro
87          180         185         190
88  Arg Arg Leu Ala Ile Gln Leu Gly Ala Gln Ala Leu Leu Gly Ala Ala
89          195         200         205
90  Lys Met Leu Leu Asp Ser Glu Gln His Pro Cys Gln Leu Lys Asp Asn
91          210         215         220
92  Val Cys Ser Pro Gly Gly Ala Thr Ile His Ala Leu His Phe Leu Glu
93          225         230         235         240
94  Ser Gly Gly Phe Arg Ser Leu Leu Ile Asn Ala Val Glu Ala Ser Cys
95          245         250         255
96  Ile Arg Thr Arg Glu Leu Gln Ser Met Ala Asp Gln Glu Lys Ile Ser
97          260         265         270
98  Pro Ala Ala Leu Lys Lys Thr Leu Leu Asp Arg Val Lys Leu Glu Ser
99          275         280         285
100 Pro Thr Val Ser Thr Leu Thr Pro Ser Ser Pro Gly Lys Leu Leu Thr
101          290         295         300
102 Arg Ser Leu Ala Leu Gly Lys Lys Asp
103          305         310
105 (2) INFORMATION FOR SEQ ID NO: 2:
107   (i) SEQUENCE CHARACTERISTICS:
108       (A) LENGTH: 1742 base pairs
109       (B) TYPE: nucleic acid
110       (C) STRANDEDNESS: single
111       (D) TOPOLOGY: linear
113   (vii) IMMEDIATE SOURCE:
114       (A) LIBRARY: PROSNON01
115       (B) CLONE: 2278458
117   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
119  GCGCCATCAG CCGCCCGGGA GATATCCGCC GGGGGAGAAT AGGGTTGCAC CATCCCAGAA 60

```

## RAW SEQUENCE LISTING

DATE: 01/02/2002

PATENT APPLICATION: US/09/912,717

TIME: 14:52:54

Input Set : N:\Crif3\RULE60\09912717.raw

Output Set: N:\CRF3\01022002\I912717.raw

```

120 GCTGCTGTTA GCTCGCCGGT CCTCGGCACG CCGCCCGTTC GCCCCTGCGC TGTCCGCCCT 120
121 TCCCCTAGCG TTACTTCCGG TCCCTCGCTG AGGGGGTTCG TGCGGCTCCC AGGAGGCGTG 180
122 AACCGCGGAC CATGAGCGTG GGCTTCATCG GGGCCGGCCA GCTGGCCTAT GCTCTGGCGC 240
123 GGGGCTTCAC GGCCGAGGC ATCCTGTCGG CTCACAAGAT AATAGCCAGC TCCCCAGAAA 300
124 TGAACCTGCC CACGGTGTCC GCGCTCAGGA AGATGGGTGT GAACCTGACA CGCAGCAACA 360
125 AGGAGACGGT GAAGCACAGC GACGTCTGTG TTCTGGCTGT GAAGCCACAT ATCATCCCCT 420
126 TCATCCTGGA TGAGATTGGG GCCGACGTGC AAGCCAGACA CATCGTGGTC TCCTGTGCGG 480
127 CTGGTGTAC CATCAGCTCT GTGGAGAAGA AGCTGATGGC ATTCCAGCCA GCCCCAAAG 540
128 TGATTTCGTG CATGACCAAC ACACCTGTGG TAGTGCAGGA AGGCGCTACA GTGTACGCCA 600
129 CGGGCACCCA TGCCCTGGTG GAGGATGGGC AGCTCCTGGA GCAGCTCATG AGCAGCGTGG 660
130 GCTTCTGCAC TGAGGTGGAA GAGGACCTCA TCGATGCCGT CACGGGGCTC AGTGGCAGCG 720
131 GGCCTGCCTA TGCATTTCAT GCTCTGGACG CATTTGGCTGA TGGTGGGGTG AAGATGGGTT 780
132 TGCCACGGCG CCTGGCAATC CAACTCGGGG CCCAGGCTTT GCTGGGAGCT GCCAAGATGC 840
133 TGCTGGACTC GGAGCAGCAT CCATGCCAGC TTAAGGACAA TGTCTGCTCC CCTGGGGGAG 900
134 CCACCATCCA CGCCCTGCAC TTTCTAGAGA GTGGGGGCTT CCGCTCTCTG CTCATCAATG 960
135 CAGTTGAGGC CTCTGTATC CGAACACGAG AGCTACAGTC CATGGCCGAC CAAGAAAAGA 1020
136 TCTCCCAGC TGCCCTTAAG AAGACCTCT TAGACAGAGT GAAGCTGGAA TCCCCACAG 1080
137 TCTCCACACT GACCCCTCC AGCCCAGGA AGCTCCTCAC AAGAAGCCTG GCCCTGGGAG 1140
138 GCAAGAAGGA CTAAGGCAGC ATCTGTCCCC TCTGTGATTC AGAGCCCTTA GTTGAGAGCC 1200
139 CCTGCCGCCC CTGCCACCCC CCTGCCCCGC TCCCACCATT GCCCTCCTC AGCTGTGCAA 1260
140 GGAGAAAGCA TGCTTAGGAA GTTTTCAGGT CCTTGTGATA AAACCTCCTT AAATCTGTTC 1320
141 AGACCAAGCA ATGCGAGCTT CCTCTCCTGT CCCATGTTGG AAGTTGCTCT GAAGGGGTGG 1380
142 TAGATGCTGG AAGCCAGACA CAACCCTGCG TACGCTGCTC AGTTGGTGA GACTGGGGCT 1440
143 GGGACTGGAG TCAGCCCAGC TGGGAGGAG GGCTGGGGAG GATCTGCAGC TGAAGCCCCG 1500
144 GGCAGGGTTG GTGTGATGCC AAGGCAAAGT GGTGAGGAGA AAACAGGAAA CGGGCTTTCT 1560
145 CTGAATTGGT AAATGGGAAA GAAGTGAGCA ACTTAAGATT GTCACAATTA ATCACAAGTG 1620
146 TACAGGATTA GACTGGGTTT ATATTTAACT CTTGCTTCAT AGGTGTACCA TTTAAAGAGT 1680
147 GTTATTTAAT GCTAAGTTA ACTGCTTTAA TAAAGTTTAT TTTTAAATAT CAAAAAATAA 1740
148 AA 1742

```

## (2) INFORMATION FOR SEQ ID NO: 3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 315 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GENBANK
- (B) CLONE: 189498

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

164 Met Ser Val Gly Phe Ile Gly Ala Gly Leu Ala Phe Ala Lys Gly
165 1 5 10 15
166 Phe Thr Ala Ala Gly Val Leu Ala Ala His Lys Ile Met Ala Ser Ser
167 20 25 30
168 Pro Asp Met Asp Leu Ala Thr Val Ser Ala Leu Arg Lys Met Gly Val
169 35 40 45
170 Lys Leu Thr Pro His Asn Lys Glu Thr Val Gln His Ser Asp Val Leu
171 50 55 60
172 Phe Leu Ala Val Lys Pro His Ile Ile Pro Phe Ile Leu Asp Glu Ile
173 65 70 75 80

```

## RAW SEQUENCE LISTING

DATE: 01/02/2002

PATENT APPLICATION: US/09/912,717

TIME: 14:52:54

Input Set : N:\Crf3\RULE60\09912717.raw

Output Set: N:\CRF3\01022002\I912717.raw

```

174 Gly Ala Asp Ile Glu Asp Arg His Ile Val Val Ser Cys Ala Ala Gly
175           85           90           95
176 Val Thr Ile Ser Ser Ile Glu Lys Lys Leu Ser Ala Phe Arg Pro Ala
177           100           105           110
178 Pro Arg Val Ile Arg Cys Met Thr Asn Thr Pro Val Val Arg Glu
179           115           120           125
180 Gly Ala Thr Val Tyr Ala Thr Gly Thr His Ala Gln Val Glu Asp Gly
181           130           135           140
182 Arg Leu Met Glu Gln Leu Leu Ser Thr Val Gly Phe Cys Thr Glu Val
183           145           150           155           160
184 Glu Glu Asp Leu Ile Asp Ala Val Thr Gly Leu Ser Gly Ser Gly Pro
185           165           170           175
186 Ala Tyr Ala Phe Thr Ala Leu Asp Ala Asp Gly Gly Val Lys Met Gly
187           180           185           190
188 Leu Pro Arg Arg Leu Ala Val Arg Leu Gly Ala Gln Ala Leu Leu Gly
189           195           200           205
190 Ala Ala Lys Met Leu Leu His Ser Glu Gln His Pro Gly Gln Leu Lys
191           210           215           220
192 Asp Asn Val Ser Ser Pro Gly Gly Ala Thr Ile His Ala Leu His Val
193           225           230           235           240
194 Leu Glu Ser Gly Gly Phe Arg Ser Leu Leu Ile Asn Ala Val Glu Ala
195           245           250           255
196 Ser Cys Ile Arg Thr Arg Glu Leu Gln Ser Met Ala Asp Gln Glu Gln
197           260           265           270
198 Val Ser Pro Ala Ala Ile Lys Lys Thr Ile Leu Asp Lys Val Lys Leu
199           275           280           285
200 Asp Ser Pro Ala Gly Thr Ala Leu Ser Pro Ser Gly His Thr Lys Leu
201           290           295           300
202 Leu Pro Arg Ser Leu Ala Pro Ala Gly Lys Asp
203           305           310           315

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/912,717

DATE: 01/02/2002

TIME: 14:52:55

Input Set : N:\Crf3\RULE60\09912717.raw

Output Set: N:\CRF3\01022002\I912717.raw

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]